

ANÁLISIS PSICOMÉTRICO DE UNA PRUEBA DE MATEMÁTICAS  
APLICADA A ESTUDIANTES DE GRADO SEXTO, UTILIZANDO LA  
TEORÍA CLÁSICA DE LOS TESTS

DEISY EDITH CLAVIJO BAQUERO  
BIBIAN ANDREA VARGAS BUSTOS

FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES  
DEPARTAMENTO DE CIENCIAS BÁSICAS  
ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA  
BOGOTÁ D.C.  
2016

ANÁLISIS PSICOMÉTRICO DE UNA PRUEBA DE MATEMÁTICAS  
APLICADA A ESTUDIANTES DE GRADO SEXTO, UTILIZANDO LA  
TEORÍA CLÁSICA DE LOS TESTS

DEISY EDITH CLAVIJO BAQUERO  
BIBIAN ANDREA VARGAS BUSTOS

Trabajo de grado realizado para obtener el título de  
Especialistas en Estadística Aplicada

Director  
WILMER DARIO PINEDA RIOS  
Magister en Matemáticas

FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES  
DEPARTAMENTO DE CIENCIAS BÁSICAS  
ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA  
BOGOTÁ D.C.  
2016

Nota de aceptación

---

---

---

---

---

---

Presidente Del Jurado

---

Jurado

---

Jurado

Bogotá, D. C. 23 de febrero de 2016

Dedicamos este triunfo a:

A Dios, por el camino recorrido....

A nuestros padres, por su amor y apoyo...

A nuestra familia, por su paciencia...

A Karina, Cristina, Brice, y Mario... amigos fieles y sinceros...

A la vida.... Por lo aprendido y aprehendido...

## AGRADECIMIENTOS

Las autoras expresan sus agradecimientos a Dios.

“Y bienaventurada la que creyó, porque se cumplirán las cosas que le  
fueron dichas de parte del Señor”  
Lucas 1: 45

Mi boca hablará sabiduría; Y el pensamiento de mi corazón inteligencia.  
Acomodaré á ejemplos mi oído: Declararé con el arpa mi enigma  
Salmos 49:3 – 4

Justicia y verdad son las obras de sus manos,  
Todos sus preceptos merecen confianza:  
Son estables para siempre jamás,  
Se han de cumplir con verdad y rectitud.  
Salmo 110; 7 - 8

## CONTENIDO

	Pág.
1.....	INT
RODUCCIÓN .....	12
1.1.....	PL
ANTEAMIENTO DEL PROBLEMA.....	14
2.....	OB
JETIVOS .....	16
2.1.....	OB
JETIVO GENERAL.....	16
2.2.....	OB
JETIVOS ESPECÍFICOS .....	16
3.....	MA
RCO DE REFERENCIA .....	17
3.1.....	LA
EVALUACIÓN PARA LA CORPORACIÓN EDUCATIVA MINUTO DE DIOS (CEMID) .....	17
3.2.....	TE
ORÍA CLÁSICA DE LOS TEST (TCT).....	17
3.2.1.....	Índi
ce de dificultad .....	18
3.2.1.1.....	Car
acterísticas .....	18
3.2.1.2.....	Inte
rpretación .....	19
3.2.2.....	Índi
ce de discriminación.....	20
3.2.2.1.....	Inte
rpretación .....	20
3.2.3.....	Con
fiabilidad .....	20
3.2.3.1.....	Inte
rpretación .....	21
3.2.4.....	Coe
ficiente de correlación biserial puntual .....	21
3.2.4.1.....	Inte
rpretación .....	22
3.2.5.....	Coe
ficiente de correlación de Pearson .....	22
3.2.5.1.....	Inte
rpretación .....	23

4.....	MA
RCO METODOLÓGICO.....	23
4.1.....	TIP
O DE ESTUDIO .....	23
4.2.....	VA
RIABLE DE ESTUDIO.....	23
4.3.....	PO
BLACIÓN Y MUESTRA.....	28
4.4.....	INS
TRUMENTOS Y MATERIALES .....	28
4.5.....	DIS
EÑO ESTADÍSTICO.....	28
5.....	AN
ÁLISIS DE RESULTADOS.....	30
5.1.....	AN
ÁLISIS DE RESULTADOS TCT .....	30
5.1.1.....	Aná
lisis de ítems .....	30
5.1.2.....	Aná
lisis de la prueba. ....	32
5.1.3.....	Aná
lisis de correlación ítem-prueba - Coeficiente de correlación biserial puntual .....	32
6.....	CO
NCLUSIONES Y RECOMENDACIONES.....	33
7.....	BIB
LIOGRAFÍA.....	35

## LISTA DE TABLAS

Tabla 1.	Dominio por ítems .....	14
Tabla 2.	Interpretación del índice de dificultad TCT.....	19
Tabla 3.	Interpretación Índice de discriminación con TCT .....	20
Tabla 4.	Interpretación Alfa de Cronbach.....	21
Tabla 5.	Interpretación del valor de Coeficiente Biserial Puntual .	22
Tabla 6.	Índices de dificultad con métodos TCT para los 20 ítems .....	30
Tabla 7.	Índices de discriminación con métodos TCT para los 20 ítems .....	31
Tabla 8.	Criterios de dificultad y discriminación por ítems. ....	31
Tabla 9.	Estadístico de confiabilidad. Alfa de Cronbach. ....	32
Tabla 10.	Índices de correlación ítem - prueba.....	33



## LISTA DE FÓRMULAS

Fórmula 1: Índice de dificultad. ....	18
Fórmula 2: Media del test. ....	19
Fórmula 3: Índice de dificultad. ....	19
Fórmula 4: Coeficiente alfa ( $\alpha$ ). TCT.....	20
Fórmula 5: Coeficiente de correlación biserial puntual. ....	21
Fórmula 6: Coeficiente de correlación de Pearson .....	21

## GLOSARIO

**ALFA DE CRONBACH ( $\alpha$ ):** es una expresión numérica que de acuerdo a una escala de medida determinar la consistencia interna del test.

**APTITUD:** capacidad y buena disposición para ejercer o desempeñar una determinada tarea, función, empleo, etc.

**CALIBRACIÓN:** proceso de ajuste de una medida hasta quedar completa y correctamente construida.

**CONFIABILIDAD:** grado de consistencia que posee el instrumento de medición.

**HABILIDAD:** corresponde al grado de competencia de un sujeto concreto frente a un objetivo determinado, a una aptitud innata o desarrollada.

**ÍNDICE DE DIFICULTAD DE UN ÍTEM:** se entiende como la proporción de personas que responden correctamente un reactivo en una prueba de aquellos que han intentado resolverlo.

**ÍNDICE DE DISCRIMINACIÓN:** determina la diferencia entre los estudiantes que tienen un rendimiento alto o bajo en un área de conocimiento.

**ÍTEM:** agente empleado en la medición. No necesariamente se refiere a preguntas de una prueba, pudiendo ser también una clasificación de un producto o cualquier otro elemento que proporcione una evidencia objetiva del objeto de medida.

**MEDICIÓN:** es un conjunto de técnicas de apoyo a la evaluación que facilitan la cuantificación objetiva de datos que sirven para el análisis y la interpretación de una situación dada.

**MUESTREO ALEATORIO SIMPLE (MAS):** técnica de muestreo en la que todos los elementos tienen idéntica probabilidad de ser seleccionados para la muestra, es un sorteo justo entre los individuos del universo.

**TEORÍA CLÁSICA DE LOS TEST (TCT):** es la teoría de los test más difundida en psicometría. El énfasis del modelo de la teoría clásica está puesto en lograr la exactitud de la medida, o en su defecto, la determinación precisa del error de medición. Por eso se le denomina con frecuencia también «teoría del error de medición». Esta teoría

pretende explicar la manera en que a partir de un valor de test medido de una persona se puede concluir el «valor verdadero» de la manifestación de la característica o rasgo de personalidad que se quiere medir.

**TEST:** Prueba de verificación que se emplea para determinar el grado de inteligencia o de capacidad en el conocimiento y aplicación de Estándares Básicos de Competencias de matemáticas.

## RESUMEN

El presente estudio determino el grado de confiabilidad que ofrece una prueba de matemáticas para grado sexto de un colegio en Soacha Cundinamarca.

Para ello se tomaron veinte ítems con respuesta de selección múltiple de un banco de preguntas de matemáticas con que cuenta la institución, los cuales corresponden a operaciones básicas con números naturales y conceptos básicos de geometría, el análisis se realizó basado en la Teoría Clásica del Test (TCT) utilizando como herramienta el software Excel y SPSS, la prueba fue aplicada a 72 estudiantes de grado sexto seleccionados mediante un muestro aleatorio simple pertenecientes al Colegio El Minuto de Dios Ciudad Verde.

Los resultados muestran que la confiabilidad del test de acuerdo al Alfa de Cronbach es de 0,73 lo que indica una consistencia interna “muy buena”, el Coeficiente de correlación biserial puntual (ítem-prueba) determina que los ítems 5, 7, 12, 13, 14, 15, 16 y 17 son consistentes con la prueba, y que al eliminar algún ítem la consistencia interna del test cambiaría.

Teniendo en cuenta los resultados del análisis, los ítems 5, 10, 13 y 15 que corresponden a operaciones básicas con Números Naturales tienen un nivel de dificultad y de discriminación adecuado de acuerdo a los rangos establecidos, por ello se recomienda no ser modificados, los demás ítems deben ser ajustados ya que no se encuentran dentro de ninguno parámetro establecido como aceptable dentro de la Teoría Clásica del Test.

## 1. INTRODUCCIÓN

“La evaluación se concibe como un proceso permanente, continuo, flexible y abierto que busca detectar en qué punto del proceso de aprendizaje va el estudiante, analiza las dificultades y limitaciones, y establece criterios para modificar y orientar procesos con el fin de generar transformaciones significativas y estructurales en los estudiantes.”<sup>1</sup>

“La evaluación se ve como una unidad de acción-reflexión-acción, que da la posibilidad de ahondar en la comprensión de los fenómenos, en la orientación que se les quiere dar y en la calidad con que se ejecutan y no como un mero ejercicio técnico para obtener resultados”<sup>2</sup>

“Al ser la evaluación un proceso, este se realiza diariamente y no siempre implica la asignación de una calificación. El éxito de esta depende en gran medida de los instrumentos utilizados para su fin como las preguntas de síntesis al final de una clase o periodo, o las preguntas de repaso sobre un tema anterior al inicio de la sesión.”<sup>3</sup>

La medición es un proceso mediante el cual se asigna valor numérico a los objetos de acuerdo con una serie de reglas establecidas, por lo que un examen debe representar fielmente el contenido y los procesos inherentes a los objetivos de aprendizaje reduciendo al mínimo las fuentes de error en la medición.<sup>4</sup>

Teniendo en cuenta las anteriores definiciones es importante que este proceso cuente con instrumentos que sean acordes al grado cognitivo de los estudiantes y que permitan validar coherentemente los procesos de aula, así alcanzar las competencias exigidas por el Ministerio de Educación Nacional - MEN y por el Sistema Institucional de Evaluación de Estudiante - SIIE.

El SIIE del Colegio El Minuto de Dios de Soacha establece que el instrumento de evaluación al finalizar cada periodo es una prueba de selección múltiple, el ICFES la define así: “Éstas le ofrecen al

---

<sup>1</sup> MINISTERIO DE EDUCACIÓN NACIONAL, Decreto 1290 de 2009, pág. 1.

<sup>2</sup> NIETO, Ligia Victoria. Manual de evaluación de desempeño. Ministerio de Educación Nacional. 2005 pág. 5

<sup>3</sup> UNIVERSIDAD EMILIO CÁRDENAS UDEC,  
<http://www.udec.edu.mx/portal/docs/DIDACTICA/INSTRUMENTOS%20DE%20EVALUACION.pdf>, Evaluación, p 8.

<sup>4</sup> Scannell, Dale. Examen y evaluación en el salón de clases. México, Editorial Diana. 1984

estudiante cuatro opciones de respuesta, de las cuales sólo una es correcta. Este tipo de pregunta puede ser utilizada para evaluar cualquiera de los cuatro procesos de lectura que son objeto de evaluación en El Estudio del Progreso Internacional en Competencia Lectora - PIRLS, por su sigla en inglés). No obstante, dado que este formato no permite que el estudiante ofrezca explicaciones que respalden su respuesta, resulta menos adecuado para evaluar la capacidad de hacer interpretaciones o evaluaciones más complejas”; es decir: de mayor complejidad para los mismos contenidos del plan curricular para grado sexto en este caso específico.<sup>5</sup>

En el presente estudio se mostrará el grado de confiabilidad que tiene una prueba con respuesta de selección múltiple sobre operaciones básicas entre números naturales y geometría básica compuesta por 20 ítems, aplicada a un grupo de 72 estudiantes de grado sexto del Colegio El Minuto de Dios Ciudad Verde de Soacha.

Para cumplir con este estudio, se analizaron los 20 ítems y la prueba en su conjunto bajo la Teoría Clásica de los Test (TCT) analizando el índice de dificultad y discriminación de los ítems, la confiabilidad de la prueba con el Alfa de Cronbach y el Coeficiente de correlación biserial puntual análisis (ítem-prueba).

El estudio es de tipo descriptivo con diseño de campo, ya que los datos se recogieron directamente de los estudiantes al aplicar el test. Los estudiantes fueron seleccionados mediante muestreo aleatorio simple utilizando para la obtención de los resultados con TCT el programa Excel y el paquete estadístico SPSS.

Cabe aclarar que el muestreo elegido fue usado en este caso ya que se tomaron los resultados de la prueba obtenidos por dos de los cuatro cursos de grado sexto de dicho colegio.

---

<sup>5</sup> INSTITUTO COLOMBIANO PARA LA EVALUACIÓN DE LA EDUCACIÓN,  
<http://www2.icfes.gov.co/examenes/saber-pro/antes-de-presentar-el-examen/instructivos-del-proceso/82-evaluaciones-internacionales/pirls/pirls/169-tipos-de-preguntas>. 2014

## 1.1. PLANTEAMIENTO DEL PROBLEMA

La Corporación educativa Minuto de Dios (CEMID) atendiendo a su misión la cual se fundamenta en el Desarrollo Humano integral con compromiso social a la luz del evangelio, busca mejorar los procesos de educación en diferentes poblaciones del país, como consecuencia de ello y con ayuda de la constructora Amarillo se fundó en el 2012 el Colegio El Minuto de Dios Ciudad Verde en Soacha (Ciudad Verde), el principal objetivo de este proyecto es ofrecer a la comunidad educación de calidad facilitando el acceso al conocimiento, al desarrollo de la inteligencia, a la formación en valores y al crecimiento en el amor por los demás<sup>6</sup>; a la institución han ingresado más de 2000 estudiantes nuevos los cuales provienen de diferentes colegios públicos o privados presentándose en consecuencia una población heterogénea en cuanto a edad y nivel educativo.

Partiendo de esto, se han venido realizando actividades de nivelación a los estudiantes de la institución con el fin de complementar y mejorar las competencias matemáticas de cada uno de ellos. En el primer periodo del año 2014 se aplicó a los estudiantes de grado 601 y 602 la evaluación bimestral cuyo objetivo es determinar el grado de habilidades en el reconocimiento y aplicación de procesos de operaciones básicas con números naturales (adición, sustracción, multiplicación, división y potenciación) y de conceptos básicos de la geometría (ángulos, punto, línea, plano, rectas paralelas y perpendiculares) en situaciones de la vida diaria; esta prueba conto con 20 ítems (ver Tabla 1.), cada uno con cuatro opciones de respuesta siendo una sola la verdadera (tipo prueba saber).

**Tabla 1. Dominio por ítems**

Ítem	Tema
1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15	Operaciones básicas con números naturales
9, 16, 17, 18, 19, 20	Conceptos básicos de la geometría

Teniendo en cuenta la escala valorativa propuesta por el SIEE de la institución el nivel de aprobación fue del 66,7 % y de reprobación del 33,3%; con preocupación se evidencia que el porcentaje de aprobación es inferior al indicador de aprobación que es del 80%.

<sup>6</sup> CORPORACIÓN EDUCATIVA MINUTO DE DIOS, <http://colegios.minutodedios.org/mision.html>, 2014

Al indagar sobre el porqué del porcentaje de reprobación y de la naturaleza de los ítems, los docentes del Departamento de matemáticas manifestaron que a la fecha de aplicación de la prueba no se hicieron estudios con antelación sobre la confiabilidad de la misma y que deducen que es poco probable que los ítems propuestos midan específicamente un contenido.

Por lo anterior el presente estudio se formula la siguiente pregunta de investigación: ¿Cuál es el grado de confiabilidad de una prueba de matemática de sexto grado a partir de la aplicación a un grupo de estudiantes del colegio El Minuto de Dios de Soacha?



## **2. OBJETIVOS**

### **2.1.OBJETIVO GENERAL**

Determinar el grado de confiabilidad y analizar algunos parámetros psicométricos que ofrece una prueba de selección múltiple aplicada a los estudiantes de grado sexto del Colegio El Minuto de Dios Ciudad verde en el I periodo de 2014 utilizando la Teoría Clásica del Test (TCT).

### **2.2.OBJETIVOS ESPECÍFICOS**

- Determinar el grado de dificultad y de discriminación de los ítems de una prueba de selección múltiple
- Establecer el grado de correlación de los ítems con la prueba aplicada
- Determinar el grado de consistencia interna y nivel de confiabilidad de la prueba haciendo uso del coeficiente Alfa de Cronbach.

### **3. MARCO DE REFERENCIA**

#### **3.1.LA EVALUACIÓN PARA LA CORPORACIÓN EDUCATIVA MINUTO DE DIOS (CEMID)**

Atendiendo al propósito de la CEMID en cuanto a proporcionar a sus estudiantes una educación pertinente, innovadora y de alta calidad, el Padre Rafael García-Herreros coincide a la evaluación como “una mirada integral que el individuo hace de sí mismo y en la que reconoce sus temores, falencias y fortalezas con la firme intención de mejorar cada día, como resultado de un compromiso consigo mismo y con la sociedad en la que habita. El ser humano está en constante evaluación: ante sí mismo, ante la sociedad y ante Dios”<sup>7</sup>.

Desde la perspectiva del MEN, la evaluación del aprendizaje de los estudiantes realizada en los establecimientos de educación básica y media, es el proceso permanente y objetivo para valorar el nivel de desempeño de los estudiantes<sup>8</sup>; así mismo define que "El alumno o educando es el centro del proceso educativo y debe participar activamente en su propia formación integral"<sup>9</sup>

Por tanto, la CEMID define la evaluación del desempeño escolar como un proceso continuo, sistemático, flexible, interpretativo, participativo y formativo que se expresa en informes cuantitativos - cualitativos - descriptivos en los cuales se emite un juicio de valor sobre el avance de los estudiantes en el alcance de los propósitos definidos en el plan de estudios<sup>10</sup>.

#### **3.2. TEORÍA CLÁSICA DE LOS TEST (TCT)**

El primer modelo en el ámbito psicológico que aborda el problema del error de las medidas realizadas mediante la aplicación de un test, fue el presentado por Spearman en 1904, donde planteó el que ya viene a ser el clásico “Modelo Lineal de Puntuaciones”, denominado “Teoría

---

<sup>7</sup> SISTEMA INSTITUCIONAL DE EVALUACIÓN DE LOS ESTUDIANTES, Gimnasio Campestre San Rafael,  
<http://colegios.minutodedios.org/SanRafael/viviendo/actividades1/Sistema%20institucional%20de%20evaluacion.pdf>, 2009-2010, pág. 1.

<sup>8</sup> MINISTERIO DE EDUCACIÓN NACIONAL, Decreto 1290 de 2009, Art. 1

<sup>9</sup> MINISTERIO DE EDUCACIÓN NACIONAL, Ley 115 de 1994, Art. 91

<sup>10</sup> CORPORACIÓN UNIVERSITARIA MINUTO DE DIOS, Documento SIE, disposiciones generales sobre la evaluación, pág. 25, 2014

débil de la Puntuaciones Verdadera” o simplemente “Teoría Clásica de los Test”.<sup>11</sup>

La Teoría Clásica de los Test (TCT) se conoce como el conjunto de principios teóricos y métodos cuantitativos derivados de ellos, que fundamentan la construcción, aplicación, validación e interpretación de distintos tipos de test y que permiten derivar escalas estandarizadas aplicables a una población los principios en que se basa son relativamente simples y se aplican tanto a las pruebas de desempeño, como a las de aptitud.

En ese sentido; en la TCT la medición del atributo se asume como la sumatoria de respuestas a ítem individuales.

### 3.2.1. Índice de dificultad

Un ítem es fácil o difícil “en función al número de personas que lo acierten o fallen en relación a l número de personas que intentan responderlo” Para este análisis basta con obtener un indicador que proporcione el índice de dificultad del ítem, el cual se obtiene de la siguiente fórmula:

**Fórmula 1:** Índice de dificultad.

$$ID = \frac{A}{N}$$

Fuente: TCT, Muñiz J, paginas 218.

Donde **A** es el número de personas que aciertan el ítem y **N** el número de personas que intentaron responderlo.

#### 3.2.1.1. Características:

Si el valor mínimo del índice de dificultad es cero “0”, se concluye que ninguno de los individuos acertó el ítem. Si el valor máximo del índice de dificultad es uno “1”, se concluye que todos los individuos acertaron el ítem. Se tiene entonces que si el índice de dificultad se acerca a cero, el ítem es difícil, si el valor se acerca a uno el ítem es fácil, y si este se aproxima a 0.5 el ítem tiene dificultad media.

---

<sup>11</sup> CHACON MOSCOSO, Salvador (2008), pág. 39. Diseño y medición de programas de intervención neuropsicológica: aspectos fundamentales.

Con base en lo anterior y de acuerdo con Guilford<sup>12</sup> retomado por López el rango aceptable para el índice de dificultad debe oscilar entre 0.2 y 0.85.

### 3.2.1.2. Interpretación:

Tabla 2. Interpretación del índice de dificultad TCT.

Valor del índice de dificultad	Criterio de clasificación
0.0 -0.15	Muy difícil
0.15 - 0.4	Difícil
0.4- 0.6	Moderado
0.6 - 0.85	Fácil
0.85-1.0	Muy fácil

El número de ítems difíciles o fáciles de una prueba depende de lo que quiere conseguir la persona que la diseña. Sin embargo en general, se aconseja que la mayoría de los ítems deban tener dificultad moderada. El valor del índice de dificultad cuando se aplica una prueba dicotómica está directamente relacionado con la media del test, es decir, la media del test es igual a la suma de los índices de dificultad de los ítems.

**Fórmula 2:** Media del test.

$$\bar{X} = \sum_{i=1}^n ID_i$$

Fuente: Muñiz, J. (1992): Teoría Clásica de los Test, Índice de discriminación, pág.: 218

Cuando los ítems son de elección múltiple y, en consecuencia, es posible acertarlos por azar, el índice de dificultad conviene calcularlo corrigiendo los afectos del azar mediante la siguiente fórmula clásica:

**Fórmula 3:** Índice de dificultad.

A: Número de sujetos que aciertan el ítem.

E: Número de sujetos que fallan el ítem.

K: Número de alternativas del ítem.

<sup>12</sup> Muñiz, J. (1992): Teoría Clásica de los Test, Guilford, J. P. (1936, 1954): Psychometric Methods, pág. 375.

N: Número de sujetos que intentan resolver el ítem.

$$ID = \frac{A - E/(K - 1)}{N}$$

Fuente: Muñiz, J. (1992): Teoría Clásica de los Test, Índice de discriminación, pág.: 219

### 3.2.2. Índice de discriminación

Se dice que un ítem tiene poder discriminativo si distingue, discrimina, entre aquellos sujetos que puntúan alto en el test y los que puntúan bajo, es decir, si discrimina entre los eficaces y los ineficaces<sup>13</sup>.

El índice de discriminación se define como la correlación entre las puntuaciones de los sujetos en el ítem y sus puntuaciones en el test. Dependiendo de las características de las variables a correlacionar, ítem-test, así se utiliza el tipo de correlación.

**Fórmula 4:** Índice de discriminación:

$A_c$ : La frecuencia de aciertos en el grupo superior (competentes).

$A_i$ : La frecuencia de aciertos en el grupo inferior (incompetentes).

$M$ : El total de individuos en cada grupo.

$$P = \frac{A_c - A_i}{M}$$

**3.2.2.1. Interpretación:** El valor del ítem depende de su valor obtenido así:

**Tabla 3.** Interpretación Índice de discriminación con TCT

Valor Obtenido	Interpretación
-1,00 a 0,00 y entre: 0.00 a 0,29	Todos los valores negativos se deben rechazar porque indicarían una discriminación contraria.
0,30 a 1,00	Sí discriminan.

### 3.2.3. Confiabilidad

<sup>13</sup> Muñiz, J. (1992): Teoría Clásica de los Test, Índice de discriminación, pág. 219.

La confiabilidad de un test es la precisión con que el test mide lo que mide, en una población determinada y en las condiciones normales de aplicación.

Es se determina utilizando el coeficiente alfa ( $\alpha$ ) propuesto por Cronbach (1951), el cual refleja el grado en el que covarían los ítems que constituyen el test, siendo entonces un indicador de la consistencia interna del test.<sup>14</sup>

**Fórmula 5:** Coeficiente alfa ( $\alpha$ ) TCT

$$\alpha = \frac{n}{n-1} \left( 1 - \frac{\sum_{j=1}^n \sigma_j^2}{\sigma_x^2} \right)$$

Fuente: Muñiz, J. (1992): Teoría Clásica de los Test, Índice de discriminación, pág.: 54

Dónde:

n: Número de ítems del test.

$\sum \sigma_j^2$ : Suma de las varianzas de los ítems.

$\sigma_x^2$ : Varianza de las puntuaciones en el test.

**3.2.3.1. Interpretación:** Dependiendo del valor de  $\alpha$ , se determina la confiabilidad de la prueba:

**Tabla 4.** Interpretación Alfa de Cronbach.

0,00 – 0,29	Baja o nula (No es confiable)
0,30 – 0,59	Confiabilidad Media
0,60 – 1	Confiabilidad Alta

### 3.2.4. Coeficiente de correlación biserial puntual ítem - prueba

Es una aplicación de la correlación de Pearson cuando una de las variables es dicotómica y la otra cuantitativa discreta o continua (la escala intervalo del puntaje de la prueba). Es el coeficiente que indica el grado de integración del ítem con la prueba. Se denomina análisis de ítem-prueba. La fórmula de la correlación biserial puntual está dada por:

**Fórmula 6:** Coeficiente de correlación biserial puntual.

<sup>14</sup> Anastasi, 1982; Aiken, 1995

$$r_{bp} = \frac{\bar{x}_p - \bar{x}}{s_x} \sqrt{\frac{p}{q}}$$

Fuente: Muñiz, J. (1992): Teoría Clásica de los Test, Índice de discriminación, pág.: 220

Dónde:

$p$ : es la proporción de individuos que acertaron.

$q$ : es la proporción de individuos que fallaron.

$\bar{x}_p$ : es la media en X de los sujetos cuya proporción es  $p$ .

$\bar{x}$ : es la media del test. y

$s_x$ : es la desviación típica del test.

#### 3.2.4.1. Interpretación

**Tabla 5.** Interpretación del valor de Coeficiente Biserial Puntual

Valor cercano a -1	El ítem tiene una baja integración
Valor cercano a 1	El ítem una alta integración con la prueba
Valor mayor a 0.30	Se acepta el ítem dentro de la prueba

## 4. MARCO METODOLÓGICO

### 4.1. TIPO DE ESTUDIO

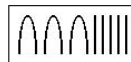
El estudio es de tipo descriptivo con diseño de campo, ya que se hace una descripción de los parámetros de discriminación, dificultad y validez los ítems bajo el modelo de la TCT.

### 4.2. VARIABLE DE ESTUDIO.

Dentro del banco de preguntas existente en la institución para ser utilizadas en las evaluaciones bimestrales, se escogieron los siguientes ítems con respuesta selección múltiple, los cuales pertenecen a operaciones básicas con números naturales y conceptos básicos de geometría.

Los ítems de la prueba aplicada fueron:

1. Un arqueólogo tiene dos documentos antiguos: uno de origen egipcio y otro de origen maya. En el primero, aparece la cantidad de sacos de trigo obtenidos en una cosecha y en el otro, la cantidad de sacos de maíz. ¿Cuál de las dos civilizaciones tuvo una mejor cosecha?



- |                                 |                        |
|---------------------------------|------------------------|
| a. Obtuvieron la misma cosecha. | c. Los egipcios.       |
| b. Los mayas.                   | d. Ninguno de los dos. |

CON LA SIGUIENTE INFORMACIÓN RESPONDE LAS PREGUNTAS 2, 3 y 4

La fábrica de juguetes “Micky” ha acelerado su nivel de producción con el fin de aumentar sus ventas en la temporada navideña. Diariamente la fábrica producirá 3750 juguetes, los cuales se empacaran en cajas grandes y medianas. En cada caja grande se acomodan 20 cajas medianas y en cada caja mediana caben 20 juguetes.

2. ¿Cuántos juguetes hay en 3 cajas grandes y 2 cajas medianas?



- a. 860 Juguetes    b. 1.100 Juguetes    c. 1.240 Juguetes    d. 1.300 Juguetes

3. ¿Cuántas cajas grandes se requieren para empacar 2.800 juguetes?

- a. 5 cajas.    b. 7 cajas.    c. 9 cajas.    d. 11 cajas.

4. El costo de la fabricación de cada juguete es \$ 4.000, si el precio de venta es \$ 10.000, ¿de cuánto es la ganancia que deja la producción diaria?

- a. \$ 37.500.000    b. \$ 22.500.000    c. \$ 15.000.000    d. \$ 7.500.000

CON LA SIGUIENTE INFORMACIÓN RESPONDE LAS PREGUNTAS 5 Y 6

SE LES PREGUNTO A LOS ESTUDIANTES DE GRADO SEXTO A QUE GRUPO DE FORMACIÓN PERTENECEN:



SE ORGANIZÓ LA INFORMACIÓN EN CONJUNTOS POR EXTENSIÓN.

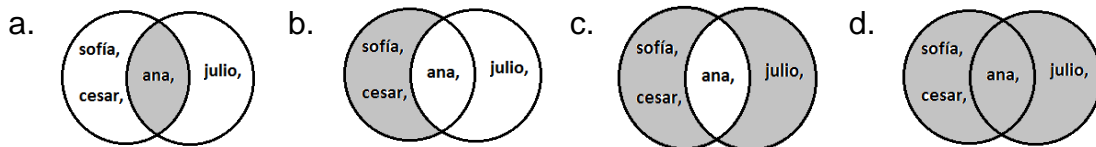
$T = \{ \text{sofía, cesar, ana} \}$

$B = \{ \text{cesar} \}$

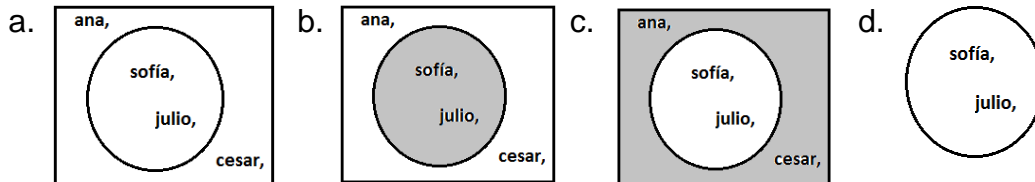
$F = \{ \text{ana, julio} \}$

$D = \{ \text{sofía, julio} \}$

5. De los siguientes diagramas de Venn el que representa correctamente  $T - F$  es:



6. Si  $U = \{ \text{sofía, cesar, ana, julio} \}$  el  $D'$  será igual a:



CON LA SIGUIENTE INFORMACIÓN RESPONDE LAS PREGUNTAS 7 Y 8

*CLAUDIA COMPRÓ VARIOS METROS DE CINTA, UNOS DE COLOR AMARILLO Y OTROS DE COLOR AZUL.*

7. Con 15 metros de cinta amarilla, Claudia puede hacer 5 adornos del mismo tamaño, iguales, sin que sobre cinta. ¿Cuántos adornos del mismo tamaño de los amarillos puede hacer con 30 metros de cinta azul sin que sobre cinta?

- a. 3      b. 5      c. 10      d. 15

8. Claudia tomó 12 metros de cinta amarilla y 20 metros de cinta azul y los cortó de forma que resultaran pedazos del mismo tamaño, no sobrara cinta y fueran de la mayor longitud posible. ¿Cuál es la longitud de cada pedazo?

- a. 3 metros      b. 4 metros.      c. 5 metros.      d. 6 metros.

CON LA SIGUIENTE INFORMACIÓN RESPONDE LAS PREGUNTAS 9 Y 10

Las boletas de entrada a un zoológico tienen un precio fijo para niños y un precio fijo para adultos. Observa el aviso que hay en la entrada del zoológico.



9. Según la información del aviso, ¿cuánto pagan 4 adultos y 6 niños por entrar en el zoológico?

- a. \$35.000      b. \$40.000      c. \$38.000      d. \$70.000

10. El precio de la boleto de un adulto es el doble del precio de la boleto de un niño. ¿Cuál es el precio de la boleto de un niño?

- a. \$5.000      b. \$20.000      c. \$7.000      d. \$25.000

11. En la siguiente tabla aparece el valor, por persona, de las boletas de entrada en un zoológico.

Días de la semana	Valor de las boletas	
	Niños y niñas	Adultos
De lunes a viernes	\$ 10.000	\$ 20.000
Sábados y domingos	\$ 12.000	\$ 25.000

Una familia compuesta por papá, mamá y tres niños entró en el zoológico el domingo.

¿Cuánto costaron las boletas de la familia?

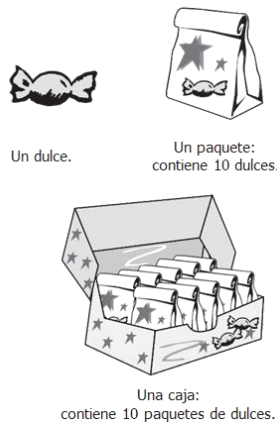
- a. \$60.000      b. \$86.000      c. \$99.000      d. \$125.000

12. El auto de Jorge necesita 6 galones de gasolina para recorrer 240 kilómetros. ¿Cuántos galones necesita para recorrer 480 kilómetros?

- a. 6      b. 8      c. 10      d. 12

CON LA SIGUIENTE INFORMACIÓN RESPONDE LAS PREGUNTAS 13 Y 14

En una dulcería se elaboraron distintos empaques para vender dulces. Observa los dibujos.



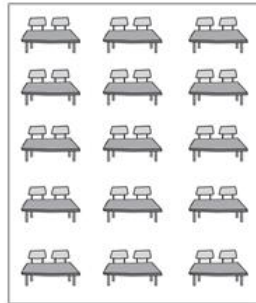
13. Doña María quiere comprar quinientos ochenta y cuatro dulces. ¿Cuántas cajas, paquetes y dulces sueltos puede comprar doña María?

- a. 4 cajas, 8 paquetes y 5 dulces sueltos.  
b. 8 cajas, 5 paquetes y 4 dulces sueltos.  
c. 5 cajas, 8 paquetes y 4 dulces sueltos.  
d. 5 cajas, 4 paquetes y 8 dulces sueltos.

14. Don Pedro compró 2 paquetes de dulces, 4 cajas de dulces y 5 dulces sueltos. ¿Cuántos dulces compró en total?

- a. 10      b. 245      c. 425      d. 542

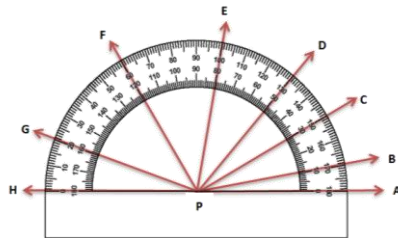
15. El siguiente dibujo muestra la organización de los pupitres dobles en un salón:



¿Con cuál de las siguientes operaciones se puede hallar el número de sillas que hay en ese salón?

- a.  $5 \cdot 3 + 2$       b.  $5 \cdot 3 \cdot 2$       c.  $(5 + 3) \cdot 2$       d.  $5 \cdot (3 + 2)$

16. El Angulo EPA tiene como medida

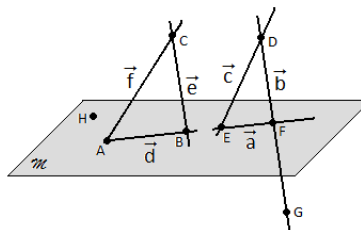


- a.  $110^\circ$     b.  $70^\circ$     c.  $80^\circ$     d.  $60^\circ$

17. El Angulo DPG tiene como medida

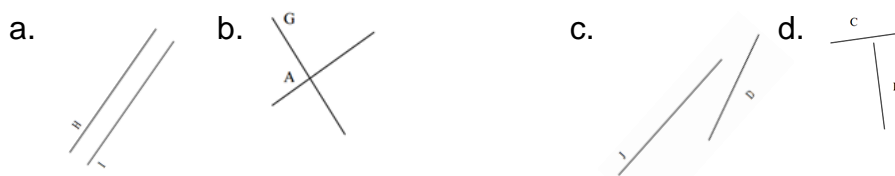
- a.  $110^\circ$       b.  $70^\circ$       c.  $80^\circ$       d.  $60^\circ$

18. En el gráfico, podemos determinar que los puntos que se encuentran en el plano M son:

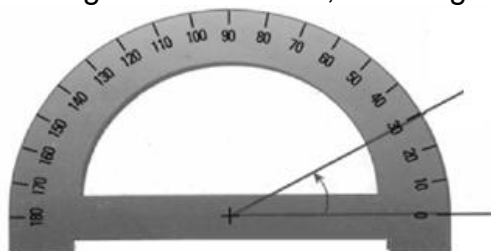


- a.  $A, B, E, F, H$     b.  $A, B, C, D, E, F, G, H$     c.  $H, G, C, D$     d.  $A, C, B$

19. De las siguientes rectas, ¿Cuáles son paralelas?



20. Según la su medida, este ángulo es:



- a. Agudo      b. Obtuso      c. Llano      d. Recto

### 4.3. POBLACIÓN Y MUESTRA

### 4.4. INSTRUMENTOS Y MATERIALES

Se seleccionaron 72 estudiantes de grado sexto del Colegio El Minuto de Dios Ciudad Verde (Soacha), seleccionados mediante un muestreo aleatorio simple.

La prueba contiene ítems sobre operaciones básicas entre números naturales (adición, sustracción, multiplicación, división y potenciación) y elementos básicos de la geometría (ángulos, punto, línea, plano, rectas paralelas y perpendiculares).

Para la validación de los datos se utilizó el software Excel y SPSS para la TCT.

### 4.5. DISEÑO ESTADÍSTICO

Para determinar la confiabilidad del test y hacer el análisis de los ítems se utilizó la Teoría Clásica del Test, para cuantificar el grado de dificultad de cada ítem se utiliza el índice de dificultad, para determinar que ítems discriminan entre estudiantes puntúan alto en el test y los que puntúan bajo se halla el índice de discriminación, para determinar el grado en que cada ítem está contribuyendo con la consistencia interna del test se calcula la correlación ítem-prueba con la correlación biserial puntual y para concluir se calcula la consistencia interna del test por medio del Alfa de Cronbach.

## 5. ANÁLISIS DE RESULTADOS.

En este apartado se hace un análisis de los ítems y de la confiabilidad de dicha prueba aplicada a los estudiantes del grado sexto del Colegio El Minuto de Dios Ciudad Verde en el mes de Marzo de 2014.

### 5.1 ANÁLISIS DE RESULTADOS TCT

#### 5.1.1. Análisis de ítems

**Tabla 6.** Índices de dificultad con métodos TCT para los 20 ítems

ítem	Dificultad	ítem	Dificultad
1	0,88**	11	0,79**
2	0,72**	12	0,61**
3	0,68**	13	0,68**
4	0,58*	14	0,65**
5	0,50*	15	0,61**
6	0,83**	16	0,57*
7	0,68**	17	0,71**
8	0,50*	18	0,44*
9	0,86**	19	0,85**
10	0,51*	20	0,72**

El rango aceptable es de 0,4 – 0,6 considerando al ítem con una dificultad moderada.

Un asterisco (\*) indica que: Se acepta el ítem

Dos asteriscos (\*\*) indican que: Se rechaza el ítem

La Tabla 6. muestra que los ítems 4, 5, 8, 10, 16, 18 están dentro del rango aceptable de dificultad en esta prueba y tienen en promedio un nivel de dificultad de 0,67; es decir la prueba tiene un nivel dificultad fácil; esto según el rango de discriminación ya establecido teóricamente para este análisis.

**Tabla 7.** Índices de discriminación con métodos TCT para los 20 ítems

Ítem	Discriminación	Ítem	Discriminación
1	0,19**	11	0,19**
2	0,28**	12	0,39*
3	0,31*	13	0,58*
4	0,28**	14	0,42*
5	0,33*	15	0,44*
6	0,17**	16	0,36*
7	0,36*	17	0,31*
8	0,28**	18	0,17**
9	0,17**	19	0,19**
10	0,36*	20	0,22**

**Discriminación:** Índice de discriminación mínimo 0,3

Un asterisco (\*)  $\geq 0,30$

indica que

Dos asteriscos (\*\*)  $< 0,30$

indican que

El ítem discrimina adecuadamente, diferencia entre los individuos de nivel superior e inferior.

El ítem no discrimina adecuadamente, no diferencia entre los individuos de nivel superior e inferior.

La tabla 7. muestra que los ítems 3, 17, 5, 7, 10, 16, 12, 14, 15 y 13 tienen un índice de discriminación mayor a 0,30, es decir, ayudan a discriminar a los estudiantes que tienen un nivel superior de los que tiene un nivel inferior.

**Tabla 8.** Criterios de dificultad y discriminación por ítems.

PARÁMETROS	CRITERIOS	VALORACIÓN	ÍTEMS
<b>DIFICULTAD</b>	0,00 a 0,15	Muy difícil	-
	0,15 a 0,40	Difícil	-
	0,40 a 0,60	Moderado	4, 5, 8, 10, 16, 18
	0,60 a 0,85	Fácil	2, 3, 6, 7, 11, 12, 13, 14, 15, 17, 19, 20
	0,85 a 1,00	Muy fácil	1, 9
<b>DISCRIMINACIÓN</b>	0,00 a 0,29	No discrimina adecuadamente	6, 9, 18, 1, 11, 19, 20, 2, 4, 8

	0,30 a 1,00	Discrimina adecuadamente	3, 17, 5, 7, 10, 16, 12, 14, 15, 13
--	-------------	-----------------------------	--

La tabla 8 clasifica a los ítems de acuerdo a su dificultad y discriminación y evidencia que:

El nivel de dificultad de los ítems “moderados” corresponde al 30% (seis ítems), del total de los ítems evaluados en la prueba, (veinte ítems).

Los ítems clasificados como “fáciles”; es decir con menor dificultad que los demás corresponden al 60% (doce ítems), del total de los evaluados en la prueba.

Las preguntas con índice de dificultad casi nula, es decir: “muy fáciles” corresponde un 10% (dos ítems) del total de las mismas.

Respecto a los ítems: 3, 5, 7, 10, 12, 13, 14, 15, 16, 17 son aceptados según su dificultad y discriminan dentro de la prueba.

Se concluye que en cuanto a la discriminación observada para toda la prueba (es decir: analizando respecto al 100% de los ítems); el 50% de ellos discriminan adecuadamente y el 50% no, es decir que 10 preguntas de las 20 realizadas muestran con los resultados obtenidos las diferencias entre los niveles superior e inferior entre los individuos evaluados y las 10 restantes no según los rangos de clasificación.

### 5.1.2. Análisis de la prueba.

**Tabla 9.** Estadístico de confiabilidad. Alfa de Cronbach.

Ítems	Alfa Cronbach
20	0,73

La consistencia interna del test es de 0,73; lo cual muestra una confiabilidad alta de acuerdo a los rangos establecidos.

### 5.1.3. Análisis de correlación ítem-prueba - Coeficiente de correlación biserial puntual



**Tabla 10.** Índices de correlación ítem - prueba.

Ítem	Correlación ítem - prueba	Ítem	Correlación ítem - prueba
1	0,204**	11	0,287**
2	0,255**	12	0,408*
3	0,262**	13	0,476*
4	0,239**	14	0,504*
5	0,383*	15	0,487*
6	0,256**	16	0,348*
7	0,341*	17	0,350*
8	0,259**	18	0,052**
9	0,158**	19	0,192**
10	0,248**	20	0,255**

Valor cercano a -1	El ítem tiene una baja integración
Valor cercano a 1	El ítem una alta integración con la prueba
Valor mayor a 0.30	Se acepta el ítem dentro de la prueba

\* >0,30 El ítem contribuye con la consistencia interna del test  
 \*\* <0,30 El ítem no contribuye con la consistencia interna del test

De acuerdo a los resultados obtenidos mediante SPSS la tabla 10 indica que los ítems 1, 2, 3, 4, 6, 8, 9, 10, 11, 18, 19 y 20 no contribuyen a la consistencia interna del test y deben ser eliminados o ajustados ya que tienen un coeficiente de correlación inferior a 0,30; los ítems 5, 7, 12, 13, 14, 15, 16 y 17 son consistentes con la prueba con un coeficiente de correlación igual o superior a 0,30.

## **6. CONCLUSIONES Y RECOMENDACIONES**

**6.1** De acuerdo nuestros objetivos planteados y a los resultados obtenidos utilizando la TCT el Índice de dificultad indica que los ítems 4, 5, 8, 10, 16, 18 están dentro del rango 0,4 - 0,6, es decir, son aceptables; al analizar el nivel de dificultad de los ítems son “moderados” el 30% (seis ítems), “fáciles” el 60% (doce ítems), y “muy fáciles” corresponde un 10% (dos ítems) del total de los ítems evaluados en la prueba (veinte ítems); así: Los ítems 4, 5, 8, 10, 16, 18 están dentro del rango 0,4 - 0,6, es decir, son aceptables.

**6.2** En cuanto a la discriminación, los ítems 3, 17, 5, 7, 10, 16, 12, 14, 15 y 13 tienen un índice mayor a 0,30, es decir, ayudan a discriminar a los estudiantes que tienen un nivel superior de los que tiene un nivel inferior.

**6.3** En cuanto al análisis del Coeficiente de correlación biserial puntual (ítem-prueba) los ítems 1, 2, 3, 4, 6, 8, 9, 10, 11, 18, 19 y 20 no contribuyen a la consistencia interna del test y deberían ser eliminados o ajustados, contrario a los ítems 5, 7, 12, 13, 14, 15, 16 y 17 que si son consistentes con la prueba teniendo en cuenta el índice de correlación, ( $>0,30$ ).

**6.4** En cuanto a la confiabilidad; utilizando el Alfa de Cronbach, la consistencia interna del test obtenida fue de 0,73; lo cual nos muestra que de acuerdo al autor y los rangos usados para comparar; el test podría tener una confiabilidad “muy buena”. Ver tabla 2, según Gilford.

**6.5** Recomendamos que para los ítems: 5, 10, 13 y 15 que corresponden a operaciones básicas con Números Naturales; como tienen un nivel de dificultad y de discriminación adecuado de acuerdo a los rangos de comparación usados; referirse a tabla 2, según Gilford y tabla 8: Criterios de dificultad y discriminación por ítems; no ser modificados. Los demás ítems deben ser ajustados ya que no se encuentran dentro de ninguno parámetro establecido como aceptable dentro de la Teoría Clásica del Test.

**6.6** Consideramos que nuestro trabajo sirvió para que los docentes de matemáticas del Colegio El Minuto de Dios Ciudad Verde fuesen conscientes de que algunos ítems diseñados para aplicar, de acuerdo a los niveles y estándares establecidos por MEN, deben ser de mayor complejidad en cuanto a lo requerido por este estamento de educación y los contenidos del plan curricular del grado sexto, en particular de la institución. Tomando en cuenta esta recomendación, se podrían establecer criterios de orientación de los procesos más apropiados hacia la temática vista.

## 7. BIBLIOGRAFÍA

BARBOSA BARBOSA, Sandra Patricia, y CALDERÓN MATEUS, Nancy Esperanza. Validez Y Confiabilidad De Un Examen De Conocimiento De Pensamiento Numérico Para Estudiantes De Pregrado. Fundación Universitaria Los Libertadores. 2013.

CABRERA, Ana Carolina, Guía No. 1. Estadística Aplicada a la Educación, Fundación Universitaria lo Libertadores.

BLANDÓN, Ana Carolina, Guía No. 2. Estadística Aplicada a la Educación, Fundación Universitaria lo Libertadores.

BLANDÓN, Ana Carolina, Guía No. 3. Estadística Aplicada a la Educación, Fundación Universitaria lo Libertadores.

BLANDÓN, Ana Carolina, Guía No. 4. Estadística Aplicada a la Educación, Fundación Universitaria lo Libertadores.

BLOG LA ACTUALIDAD SOBRE LA INVESTIGACIÓN POR INTERNET, <http://www.netquest.com/blog/es/muestreo-probabilistico-muestreo-aleatorio-simple/>

CHACÓN MOSCOSO, Salvador (2008), pág. 39. Diseño y medición de programas de intervención neuropsicológica: aspectos fundamentales.

CORPORACIÓN EDUCATIVA MINUTO DE DIOS, disponible es <<http://colegios.minutodedios.org/mision.html>>, 2014

CORPORACIÓN UNIVERSITARIA MINUTO DE DIOS, Documento SIIE 2014.

GIMNASIO CAMPESTRE SAN RAFAEL, Sistema Institucional De Evaluación De Los Estudiantes, <http://colegios.minutodedios.org/SanRafael/viviendo/actividades1/Sistema%20institucional%20de%20evaluacion.pdf>.

GUILFORD, J. P. (Psychometric Methods, 1936, 1954.

INSTITUTO COLOMBIANO PARA LA EVALUACIÓN DE LA EDUCACIÓN, disponible en <<http://www2.icfes.gov.co/examen/saber-pro/antes-de-presentar-el-examen/instructivos-del-proceso/82-evaluaciones-internacionales/pirls/pirls/169-tipos-de-preguntas>>. 2014

MINISTERIO DE EDUCACIÓN NACIONAL, Decreto 1290 de 2009.

MINISTERIO DE EDUCACIÓN NACIONAL, Ley 115 de 1994.

Muñiz, José. Teoría Clásica de los Test, 1992

NIETO, Ligia Victoria. Manual de evaluación de desempeño. Ministerio de Educación Nacional.

SCANNELL, Dale. Examen y evaluación en el salón de clases. México, Editorial Diana. 1984

UNIVERSIDAD EMILIO CÁRDENAS UDEC, Evaluación, disponible en <<http://www.udec.edu.mx/portal/docs/DIDACTICA/INSTRUMENTOS%20DE%20EVALUACION.pdf>>.